



DEMIG TOTAL Database Codebook

This document provides information on DEMIG TOTAL database, its worksheets and their layout. The data file contains four worksheets (tabs):

- Summary
- Total Inflows
- Total Outflows
- Net Flows

Following are the descriptions of the contents of these worksheets.

~~~~~

The **Summary** provides an overview of the data in this order:

- **Number:** sequential country number in the database
- **Country:** name of country for which we report total migration flows
- **Major area:** each country is classified according to the major area to which it belongs (follows UN-adopted categorisations)
- **Region:** each country is classified according to its region (e.g. Caribbean, South East Asia, etc., also following UN-adopted categorisations)
- **Data origin:** identifies whether the data were entered manually from tables (T), they were electronic data (E) or a combination of tables and electronic (M)
- **Type of sources:** provides information on the *data collection method* utilised; unfortunately this information is not always available; as a result for many countries this is completed as 'Unknown.'
- **Inflows (years):** lists the years for which data are reported; some years are followed by short notes between parentheses to make clear to the user that there were changes in types of flows reported (e.g. short-term, citizens only, etc.).
- **C/F/T:** these cells are to indicate whether the data are reported for citizens, foreigners and/or totals.
- **Source:** lists the sources for inflow data; the sources follow the chronological order of the series to guide any eventual retracing.
- **Outflow (years):** lists the years for which data are reported; some years are followed by short notes between parentheses to make clear to the user that there were changes in types of flows reported (e.g. short-term, citizens only, etc.).
- **C/F/T:** these cells are to indicate whether the data are reported for citizens, foreigners and/or totals.

- **Source:** lists the sources for outflow data; the sources follow the chronological order of the series to guide any eventual retracing.
- **Net flows (years):** lists the years for which data are reported; some years are followed by short notes between parentheses to make clear to the user that there were changes in types of flows reported (e.g. short-term, citizens only, etc.).
- **C/F/T:** these cells are to indicate whether the data are reported for citizens, foreigners and/or totals.
- **Source:** this lists the sources for the net flows; the sources follow the chronological order of the series to guide any eventual retracing.
- **Notes:** in these cells you find notes for the entire series and details on the nature of the flows, some indication to collection methods, rounded figures, etc.
- **Historical background:** this information originates almost exclusively from Mitchell 2005, which has very useful historical facts and information on boundary changes. These have been entered as a useful additional reference.

~~~~~

The **Total Inflows**, **Total Outflows** and **Net Flows** worksheets are data sheets. The data cells report the international moves recorded in each specific year for particular categories as defined by the columns. The three worksheets share the same layout, with 11 coding columns before the data columns, which covers 196 years in the 1815-2011 period.

The columns code the data according to these criteria:

- **Type of flow:** Inflows, Outflows or Net flows
- **Coverage:** Citizens/Foreigners/Total
- **Country:** name of country for which migration flows are reported, i.e. the row reporting inflows, foreigners, Australia displays how many foreigners entered Australia in each specific year.
- **UN numeric code:** this field matches the 'Country' with a unique numeric code based on the UN coding system.
- **Country codes-UN based:** this field matches the 'Country' with a unique alphabetic code based on the UN coding system.
- **Country codes-WDI based:** this field matches the 'Country' with a unique alphabetic code based on the coding system used in the World Development Indicators.
- **Major area:** each country is classified according to Major areas (continents + former USSR).
- **Major area code:** this field is coded on the basis of the UN coding system.
- **Region:** each country is classified according to its region (e.g. Caribbean, South East Asia, etc. + former USSR)

- **Region code:** this field is coded on the basis of the UN coding system.
- **Year:** 1815-2011

Colouring and comments in values: Values are generally reported in black. However, in some cases, the longitudinal series are interjected with data that originate from different sources – when that is the case the interjecting data have been written in purple.

Figures may also be highlighted in yellow. This is because the definition has changed over the years or the data are questionable. In the 1950-late 1970s you may find a number of figures in yellow highlighting, because the source switched from reporting flows for long-term immigrants and emigrants to figure reporting together long-term and short-term migration. However, these figures had some inconsistency.

Often figures in purple or yellow are accompanied by comments linked to the cells with specific information about that cell (or series of cells). Cells with figures in black may also be accompanied by comments to draw the attention of the user to data specificities.

Note on codes: The UN codes are based on lists and codes provided by the United Nations in 2010. For countries that no longer exist (e.g. Czechoslovakia, German Democratic Republic), we used UN codes that have been 'retired'. The WDI codes adopted the coding system utilised by the World Development Indicators (WDI). The WDI coding system does not make available codes for countries that no longer exist; therefore, the same 'retired' codes provided by the UN were adopted.